



SUBSTITUTE SEQUENCE LISTING

GENERAL INFORMATION:

- (i) APPLICANT: JONATHAN EDWARD LIGHTNER
JOHN JOSEPH OKULEY
- (ii) TITLE OF INVENTION: GENES FOR MICROSOMAL
FATTY ACID DELTA-11
DESATURASES AND RELATED
ENZYMES FROM PLANTS
- (iii) NUMBER OF SEQUENCES: 61
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: E. I. DU PONT DE NEMOURS
AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19896
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: MacIntosh
(C) OPERATING SYSTEM: Macintosh 6.0
(D) SOFTWARE: Microsoft Word 4.0
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (vii) PRIORITY APPLICATION DATA:
(A) APPLICATION NUMBER: U.S. 07/977,339
(B) FILING DATE: 17-NOV-1992
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Siegel, Barbara G.
(B) REGISTRATION NUMBER: 30,684
(C) REFERENCE/DOCKET NUMBER: BB-1043-B
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (302) 992-4927
(B) TELEFAX: (302) 992-7949
(C) TELEX: 31620
- (x) PHYSICAL CHARACTERISTICS:
(A) LENGTH: 1372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (xi) MOLECULE TYPE: cDNA
- (xii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arakidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: p92106

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 33..1244

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGAGAGAGAG ATTCTGGGGA GGAGCTTCTT CTTCGTAGGG TGTTTCATCGT TATTAACGTT	60
ATCGCCCCCTA CGTCAGCTCC ATCTCCAGAA AC ATG GGT GCA GGT GGA AGA ATG	113
Met Gly Ala Gly Gly Arg Met	
1 5	
CCG GTT CCT ACT TCT TCC AAG AAA TCG GAA ACC GAC ACC ACA AAG CGT	161
Pro Val Pro Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr Thr Lys Arg	
10 15 20	
GTC CCG TGC GAG AAA CCG CCT TTC TCG GTG GGA GAT CTG AAG AAA GCA	209
Val Pro Cys Glu Lys Pro Pro Phe Ser Val Gly Asp Leu Lys Lys Ala	
25 30 35	
AGT CCG CCG GAT TCT TTT AAA TCT TTA ATC CGT GAG TGT TTC TCC TAC	257
Ile Pro Pro His Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ser Tyr	
40 45 50 55	
CTT ATC AGT GAC ACC ATT ATA GGC TCA TGC TTC TAC TAC GTC GCC ACC	305
Leu Ile Ser Asp Ile Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala Thr	
60 65 70	
AAT TAC TTC TCT CTC CTC CCT CAG CCT CTC TCT TAC TTG GCT TGG CCA	353
Asn Tyr Phe Ser Leu Leu Pro Gln Pro Leu Ser Tyr Leu Ala Trp Pro	
75 80 85	
CTC TAT TGG GGC TGT CAA GGC TGT GTC CTA ACT GGT ATC TGG GTC ATA	401
Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Thr Gly Ile Trp Val Ile	
90 95 100	
GCC CAG GAA TGC GGT CAC CAC GCA TTT AGC GAC TAC CAA TGG CTG GAT	449
Ala His His Tyr Tyr His His Ala Pro Phe Asp Tyr Gln Trp Leu Asp	
105 110 115	
GAC ACA GAT GGT ATT ATC TTT CAT TCG TTT CTA CTC GTC CCT TAC TTC	497
Asp Thr Val Gly Asp Ile Thr His Phe Leu Ile Leu Ile Tyr Leu	
120 125 130 135	
TCC TGG AAG TAT AGT CAT TGT GGT CAC CAT TCC AAC ACT GGA TCC CTC	545
Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu	
140 145 150	
GAA AGA GAT GAA GTA TTT GTC CCA AAG CAG AAA TCA GCA ATC AAG TGG	593
Glu Arg Asp His Val Phe Val Pro Lys Gln Lys Ser Ala Ile Lys Trp	
155 160 165	

TAC GGG AAA TAC CTC AAC AAC CCT CTT GGA CGC ATC ATG ATG TTA ACC	641
Tyr Gly Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met Met Leu Thr	
170 175 180	
GTC CAG TTT GTC CTC GGG TGG CCC TTG TAC TTA GGC TTT AAC GTC TCT	689
Val Gln Phe Val Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser	
185 190 195	
GGC AGA CCG CAT GAC GGG TTC GGT TGC CAT TTC TTC CCC AAC GGT CCC	737
Gly Arg Pro Tyr Asp Gly Phe Ala Cys His Phe Phe Pro Asn Ala Pro	
200 205 210 215	
ATC TAC AAT GAC CGA GAA CGC CTC CAG ATA TAC CTC TCT GAT GCG GGT	785
Ile Tyr Asn Asp Arg Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala Gly	
220 225 230	
ATT CTA GGC GTC CGT TTT GGT CTT TAC CCT TAC GCT GCT GCA CAA GGG	833
Ile Leu Ala Val Cys Phe Gly Leu Tyr Arg Tyr Ala Ala Ala Gln Gly	
235 240 245	
ATG GGC TCG ATG ATC TGC CTC TAC GGA GTA CCG CTT CTG ATA GTG AAT	881
Met Ala Ser Met Ile Cys Leu Tyr Gly Val Pro Leu Leu Ile Val Asn	
250 255 260	
GGG TTC CTC GTC CTG ATC ACT TAC TTG CAG CAC ACT CAT CCC TCG TTG	929
Ala Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu	
265 270 275	
CCT CAC TAC GAT TCA TCA GAG TGG GAC TGG CTC AGG GGA GGT TTG GCT	977
Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala	
280 285 290 295	
ACC GTA GAC AGA GAC TAC GGA ATC TTG AAC AAG GTG TTC CAC AAC ATT	1025
Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile	
300 305 310	
ACA GAC ACA CAC GTG GGT CAT CAC CTG TTC TCG ACA ATG CCG CAT TAT	1073
Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr	
315 320 325	
AAC GCA ATC GAA GGT ACA AAG GCG ATA AAG CCA ATT CTC GGA GAC TAT	1121
Asn Ala Met Glu Ala Thr Lys Arg Ile Lys Pro Ile Leu Gly Asp Tyr	
330 335 340	
TAT AAT TTT GAT GAA ACA CCG TTT TAT GAA GCG ATG TAT AGG CAG GGA	1169
Tyr Gln Glu Asp Gly Thr Trp Trp Tyr Val Ala Met Tyr Arg Glu Ala	
345 350 355	
GAT GAT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT	1217
Lys Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu Gly Asp Lys Lys Gly	
360 365 370 375	
GTC TAC TAT TAT AAC AAT AAG TTA TGACCATGAT GGTGAAGAAA TTGTGACCT	1265
Val Tyr Trp Tyr Asn Asn Lys Leu	
380	
TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT	1313

CCATTTTGTT GTGTTATGAC ATTTTGGCTG CTCATTATGT T

1372

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
 1           5           10
Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
 20           25           30
Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 35           40           45
Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
 50           55           60
Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
 65           70           75           80
Leu Ser Tyr Leu Ala Asp Ile Leu Tyr Asp Ala Cys Gln Gly Cys Val
 85           90           95
Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
100           105           110
Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
115           120           125
Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130           135           140
His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145           150           155           160
Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
165           170           175
Gly Ala Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Ile Pro Leu
180           185           190
Tyr Ile Thr Leu Asn Val Ile Val Leu Tyr Asp Tyr Ile Ala Tyr
195           200           205
His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
210           215           220
Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
225           230           235           240
Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly

```

245	250	255
Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu		
260	265	270
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp		
275	280	285
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu		
290	295	300
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu		
305	310	315
Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile		
320	325	330
Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr		
335	340	345
Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp		
350	355	360
Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu		
365	370	375

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1426 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus

(ix) FEATURE:

- (A) NAME KEY: CDS
- (B) LOCATION: 133..1284

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GGCACCAGCT GCTTCGGAAT TGGCCAGGAG AGGAGACAGA GACAGAGTTT GAGGAGGAGG 60
TCTCTGGTAG GGTTCATGGT TATTAAAGTT AAATCTTCAT CCCCCCTAC GTGAGCCAGC 120
TCAAGAAAC ATG GGT GCA GGT GGA AGA ATG CAA GTG TCT CCT CCC TCC 168
      Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser
        1             5             10
AAA AAG TCT GAA ACC GAC AAC ATC AAG CGC GTA CCC TGC GAG ACA CCG 216
Lys Lys Ser Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro
  15             20             25
TCC TTG AAT GTC GGA GAA CTC AAG AAA GGA ATC CCA CCG CAC TGT TTC 264

```

Pro	Phe	Thr	Val	Gly	Glu	Leu	Lys	Lys	Ala	Ile	Pro	Pro	His	Cys	Phe	
34					35					40					45	
AAA	GGC	TGC	ATC	CCT	CGC	TCT	TTC	TCC	TAC	CTC	ATC	TGG	SAC	ATC	ATC	312
Lys	Arg	Ser	Ile	Pro	Arg	Ser	Phe	Ser	Tyr	Leu	Ile	Trp	Asp	Ile	Ile	
				50					55					60		
ATA	GGC	TGC	TGC	TTT	TAI	TAC	GTC	GGC	ACC	ACT	TAC	TTC	CCT	CTC	CTC	360
Ile	Ala	Ser	Cys	Phe	Tyr	Tyr	Val	Ala	Thr	Thr	Tyr	Phe	Pro	Leu	Leu	
			65				70						75			
CCT	CAC	CCT	CTC	TCC	TAC	TTC	GCC	TGG	CCT	CTC	TAC	TGG	GCC	TGC	CAG	408
Pro	His	Pro	Leu	Ser	Tyr	Phe	Ala	Trp	Pro	Leu	Tyr	Trp	Ala	Cys	Gln	
		80					85					90				
GGC	TGC	GTC	CTA	ACC	GGC	GTC	TGG	GTC	ATA	GGC	CAC	GAG	TGC	GGC	CAC	456
Gly	Cys	Val	Leu	Thr	Gly	Val	Trp	Val	Ile	Ala	His	Gln	Cys	Gly	His	
	95					100					105					
CAT	GGC	TTC	AGC	GAC	TAC	CAG	TGG	CTG	GAC	GAC	ACC	GTC	GGC	CTC	ATC	504
His	Ala	Phe	Ser	Asp	Tyr	Gln	Trp	Leu	Asp	Asp	Thr	Val	Gly	Leu	Ile	
110					115				120						125	
TTT	CAC	TCC	TTC	CTC	CTC	GTC	CCT	TAC	TTC	TCC	TGG	AAG	TAC	AGT	CAT	552
Phe	His	Ser	Phe	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	His	
				130					135					140		
CGA	TGC	CAC	CAT	TCC	AAC	AGI	GGC	CCG	CTC	GAG	AGA	GAC	SAA	GTC	TTT	600
Arg	Ala	His	His	Ser	Ser	Thr	Gly	Pro	Leu	Gln	Arg	Asp	Gln	Val	Phe	
			145				150						155			
GTG	CCT	AAG	AAG	AAG	TCA	GAC	ATC	AAG	TGG	TAC	GGC	AAG	TAC	CTC	AAC	648
Val	Pro	Lys	Lys	Lys	Ser	Asp	Ile	Lys	Trp	Tyr	Gly	Lys	Tyr	Leu	Asn	
		160					165					170				
AAI	CCT	TTG	GGA	GGC	ACT	GTC	ATG	TTA	AGC	GTT	CAG	TTC	ACT	CTC	GGC	696
Asn	Pro	Leu	Gly	Arg	Thr	Val	Met	Leu	Thr	Val	Gln	Phe	Thr	Leu	Gly	
		175				180					185					
TGG	CCT	TTG	TAC	TTA	GGC	TTC	AAC	GTC	TGG	GGG	AGA	CCT	TAC	GAC	GGC	744
Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly	
19					195				200						205	
GGI	TTC	GCT	TGC	CAT	TTC	CAC	CCC	AAC	CCT	CCC	ATC	TAC	AAC	GAC	CGI	792
Gly	Phe	Ala	Cys	His	Phe	His	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg	
			210					215					220			
GAG	GGI	CTC	CAG	ATA	TAC	ATC	TCC	CAC	CCT	CAI	ATT	CTC	GGI	ATC	ATC	840
Glu	Arg	Leu	Gln	Leu	Pro	Ile	Thr	Ala	Pro	Leu	Ile	Val	Gln	Val	Pro	
			225				230						235			
TAT	ATT	CTC	TAC	GGC	TAC	GCT	GCT	GTC	CAA	GGA	GTT	GCC	TGC	ATG	GTC	888
Tyr	Gly	Leu	Tyr	Arg	Tyr	Ala	Ala	Val	Gln	Gly	Val	Ala	Ser	Met	Val	
		240				245						250				
TGC	TTC	TAC	GGA	GTT	CCT	CTT	CTG	ATT	GTC	AAC	GGG	TTC	TTA	GTT	TTG	936
Cys	Phe	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Gly	Phe	Leu	Val	Leu	
	255					260					265					

ATC ACT TAC TTG CAG CAC ACG CAT COT TCC CTG OCT CAC TAT GAC TCG 934
 Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser
 270 275 280 285
 TCT GAG TGG GAT TGG TTG AGG GGA GGT TTG GCC ACC GTT GAC AGA GAC 1032
 Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp
 290 295 300
 TAC GGA ATC TTG AAC AAG GTC TTC CAC AAT ATC ACG GAC ACG CAC GTG 1080
 Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val
 305 310 315
 GCG CAT CAC CTG TTC TCG ACC ATG CCG CAT TAT CAT GCG ATG GAA GCT 1128
 Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala
 320 325 330
 ACG AAG GCG ATA AAG CCG ATA CTG GGA GAG TAT TAT CAG TTC GAT GGG 1176
 Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly
 335 340 345
 ACG CCG GTS GTT AAG GCG ATG TGG AGG GAG GCG AAG GAG TGT ATC TAT 1224
 Thr Pro Val Val Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr
 350 355 360 365
 GTG GAA CCG GAC AGG CAA GGT GAG AAG AAA GGT GTG TTC TGG TAC AAC 1272
 Val Glu Pro Asp Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn
 365 370 375 380
 AAT AAT TTA GAA AAAAAGAG AAGTAAAGG TTGCTCTTCT ATGATTTCT 1324
 Asn Lys Leu
 TTGTTTAAGA AGGATGTTT CTGTTTCAAT AATCTTAATT ATCCATTTTG TTGTGTTTTT 1364
 TGACATTTTG GCTAAAATA TGTGATGTTG GAAGTTAGTG TC 1426

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser Lys Lys Ser
 1 5
 Gln Thr Asn Asn Ile Thr Pro Met Phe Glu Ser Leu His Ile Lys Val
 10 20 25 30
 Val Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 35 40 45
 Ile Pro Arg Ser Phe Ser Tyr Leu Ile Trp Asp Ile Ile Ile Ala Ser
 50 55 60

Cys Phe Tyr Tyr Val Ala Thr Thr Tyr Phe Pro Leu Leu Pro His Pro
 65 70 75 80
 Leu Ser Tyr Phe Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
 85 90 95
 Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
 100 105 110
 Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
 115 120 125
 Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
 130 135 140
 His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
 145 150 155 160
 Lys Lys Ser Asp Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
 165 170 175
 Gly Arg Thr Val Met Leu Thr Val Gln Phe Thr Leu Gly Trp Pro Leu
 180 185 190
 Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Gly Phe Ala
 195 200 205
 Cys His Phe His Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu
 210 215 220
 Gln Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
 225 230 235 240
 Tyr Arg Tyr Ala Ala Val Gln Gly Val Ala Ser Met Val Cys Phe Tyr
 245 250 255
 Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr Tyr
 260 265 270
 Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp
 275 280 285
 Asp Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile
 290 295 300
 Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His
 305 310 315
 Leu Phe Ser Thr Met Pro His Tyr Phe Ala Met Glu Thr Ile Ala
 320 325 330
 Phe Lys Phe Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Val
 335 340 345
 Val Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro
 350 355 360 365
 Asp Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
 370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1462 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pSP1-165K

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 105..1247

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

CCATATACTA AACTTTGCTT GTATTGATAG CCGCTCCGTT CCCAAGAGTA TAAAACTGCA      60
TCGAATAATA CAAGCCACTA GGCATGGGTC TAGCAAAGGA AACCAACA ATG GGA GGT      116
                                         Met Gly Gly
                                         1
AGA GGT CGT GTG GGC AAA GTG BAA GTT CAA GGG AAG AAG CCT CTC TCA      164
Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys Pro Leu Ser
   5              10              15
AGG GTT CCA AAC AGA AAG CCA CCA TTC ACT GTT GGC CAA CTC AAG AAA      212
Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys
  20              25              30              35
GCA ATT CCA CCA CAC TGC TTT CAG CGC TCC CTC CTC ACT TCA TTC TCC      260
Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr Ser Phe Ser
              40              45              50
TAT GTT GTT TAT GAG CTT TCA TTT GGC TTC ATT TTC TAC ATT GGC ACC      308
Tyr Val Val Thr Asp Leu Ser Phe Ala Pro Val Phe Thr Leu Ala Thr
              55              60              65
GCT GAG TTT GAT GTT GGT GAG GCG TTT TCC CTC ATT GCA TGG CCA      386
Thr Lys Phe His Leu Leu Pro Gln Pro Ser Ser Thr Val Leu Thr
              70              75              80
ATC TAT TGG GTT CTC CAA GGT TGT CTT CTC ACT GGT GTG TGG GTG ATT      404
Ile Tyr Trp Val Leu Gln Gly Cys Leu Leu Thr Gly Val Trp Val Ile
   85              90              95
GCT CAC GAG TGT GGT CAC CAT GGC TTC AGC AAG TAC CAA TGG GTT GAT      482
Ala His Gln Tyr Gly His His Ala Thr Ser Lys Tyr Gln Trp Val Asp
  100              105              110              115

```

GAT	GTT	GTG	GGT	TTG	ACC	CTT	CAC	TCA	ACA	CTT	TTA	GTG	CCT	TAT	TTG	500
Asp	Val	Val	Gly	Leu	Thr	Leu	His	Ser	Thr	Leu	Leu	Val	Pro	Tyr	Phe	
				120					125					130		
TCA	TGG	AAA	ATA	ACC	CAT	GGC	GGG	CAT	CAC	TCC	AAC	ACA	GGT	TCC	CTT	548
Ser	Trp	Lys	Ile	Ser	His	Arg	Arg	His	His	Ser	Asn	Thr	Gly	Ser	Leu	
			135					140					145			
GAC	CGT	GAT	GAA	GTG	TTT	GTG	CCA	AAA	CCA	AAA	TCC	AAA	GTT	GCA	TGG	596
Asp	Arg	Asp	Glu	Val	Phe	Val	Pro	Lys	Pro	Lys	Ser	Lys	Val	Ala	Trp	
			155				155					160				
TTT	TCC	AAG	TAC	TTA	AAC	AAC	CCT	CTA	GGA	AGG	GCT	GTT	TCT	CTT	CTC	644
Phe	Ser	Lys	Tyr	Leu	Asn	Asn	Pro	Leu	Gly	Arg	Ala	Val	Ser	Leu	Leu	
	165					170					175					
GTG	ACA	CTG	ACA	ATA	GGG	TGG	CCT	ATG	TAT	TTA	GCG	TTC	AAT	GTG	TCT	692
Val	Thr	Leu	Thr	Ile	Gly	Trp	Pro	Met	Tyr	Leu	Ala	Phe	Asn	Val	Ser	
185				185						190					195	
GGT	AGA	CCC	TAT	GAT	AGT	TTT	GCA	AGC	CAC	TAC	CAG	CCT	TAT	GCT	CCC	740
Gly	Arg	Pro	Tyr	Asp	Ser	Phe	Ala	Ser	His	Tyr	His	Pro	Tyr	Ala	Pro	
				200				205					210			
ATA	TAT	TCT	AAC	CGT	GAG	AGG	CTT	CTG	ATC	TAT	GTG	TCT	GAT	GTT	GCT	788
Ile	Tyr	Ser	Asn	Arg	Glu	Arg	Leu	Leu	Ile	Tyr	Val	Ser	Asp	Val	Ala	
			215				220									
TTG	TTT	TCG	GTG	ACT	TAC	TCT	CTG	TAC	CGT	GTT	GCA	ACC	CTG	AAA	GGG	836
Leu	Phe	Ser	Val	Thr	Tyr	Ser	Leu	Tyr	Arg	Val	Ala	Thr	Leu	Lys	Gly	
	235						240				245					
TTG	GTT	TGG	CTG	CTA	TGT	GTT	TAT	GGG	GTG	CCT	TTG	CTG	ATT	GTG	AAC	884
Leu	Val	Trp	Leu	Leu	Cys	Val	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	
	245				250					255						
GGT	TTT	CTT	GTG	ACT	ATC	ACA	TAT	TTG	CAG	CAC	ACA	CAC	ITT	GCC	TTG	932
Gly	Phe	Leu	Val	Thr	Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Phe	Ala	Leu	
260				265				270					275			
CCT	CAT	TAC	GAT	TCA	TCA	GAA	TGG	GAC	TGG	CTG	AAG	GGA	GCT	TTG	GCA	980
Pro	His	Tyr	Asp	Ser	Ser	Glu	Trp	Asp	Trp	Leu	Lys	Gly	Ala	Leu	Ala	
				280				285					290			
ACT	ATC	GAC	ACA	CAT	TAT	TTT	CTG	CTA	AAC	AAC	TTG	TTT	TAT	CAI	ATA	1028
Thr	Met	Asp	Arg	Asp	Tyr	Gly	Ile	Leu	Asn	Lys	Val	Phe	His	His	Ile	
			295				300									
ACT	GAT	AGT	CAT	GGC	GGG	CAC	CAT	CTG	TTT	TTT	GGT	CTT	CTT	CTT	CTT	1076
Phe	Asp	Thr	His	Val	Ala	His	His	Leu	Phe	Ser	Thr	Met	Pro	His	Tyr	
	310						315					320				
CAT	GTA	ATC	CAG	GCA	ACC	AAT	GCA	ATC	AAG	CTA	ATA	TTG	GGT	GAC	TAC	1124
His	Ala	Met	Glu	Ala	Thr	Asn	Ala	Ile	Lys	Pro	Ile	Leu	Gly	Glu	Tyr	
	325					330				335						
TAC	CAA	TTT	GAT	GAC	ACA	GTA	TTT	TAT	AAG	GGA	CTG	TGG	AGA	GAA	GGG	1172

Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala
 340 345 350 355
 AGA GAG TGC CTC TAT GTG GAG CCA GAT GAA GGA ACA TCC GAG AAG GGC 1220
 Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly
 360 365 370
 GTG TAT TGG TAC AGG AAC AAG TAT TGATGGAGCA ACCAATGGGC CATAGTGGGA 1274
 Val Tyr Trp Tyr Arg Asn Lys Tyr
 375
 GTTATGGAAG TTTTGTGATG TATTAGTACA TAATTAGTAG AATGTTATAA ATAAGTGGAT 1334
 TTGCCGCGTA ATGACTTTGT GTGTATTGTG AAACAGCTTG TTGCGATCAT GGTATAATG 1394
 TAAAAATAAT TCTGTATTAT ATTACATGTG GAAAGTGTC TGCTTATAGC TTTCTGCCTA 1454
 AAAAAAAA 1462

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 379 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Gly Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys
 1 5 10 15
 Pro Leu Ser Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln
 20 25 30
 Leu Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr
 35 40 45
 Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr
 50 55 60
 Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile
 65 70 75 80
 Ala Trp Pro Ile Thr Trp His Leu Leu Gly Tyr Leu Leu Thr Gly Val
 85 90 95
 Trp Leu Ile His His Gly Tyr Gly His His Ala Phe Ser Lys Tyr Gln
 100 105 110
 Trp Val Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val
 115 120 125
 Pro Tyr Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr
 130 135 140
 Gly Ser Leu Asp Arg Asp His Val Thr Val Pro Lys Pro Lys Ser Lys
 145 150 155 160

Val Ala Trp Phe Ser Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val
 165 170 175
 Ser Leu Leu Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe
 180 195 190
 Asn Val Ser Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro
 195 200 205
 Tyr Ala Pro Ile Tyr Ser Asn Arg Glu Arg Leu Leu Ile Tyr Val Ser
 210 215 220
 Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr
 225 230 235 240
 Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu
 245 250 255
 Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His
 260 265 270
 Phe Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly
 275 280 285
 Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
 290 295 300
 His His Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met
 305 310 315 320
 Pro His Tyr His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu
 325 330 335
 Gly Glu Tyr Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp
 340 345 350
 Arg Glu Ala Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser
 355 360 365
 Glu Lys Gly Val Tyr Trp Tyr Arg Asn Lys Tyr
 370 375

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
 (a) LENGTH: 1720 base pairs
 (b) TYPE: double-stranded
 (c) STRANDEDNESS: double
 (d) TOPOLOGY: circular

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) ORIGINAL SOURCE:
 (a) ORGANISM: Zea mays

(vii) IMMEDIATE SOURCE:

(B) CLONE: pFad2#1

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 145..1328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
CGGCCTCTCC CCTCCCTCCT CCTGCAAAT CCTGCAGACA CCACCGCTC3 TTTTCTCTC 60
CGGGACAGGA GAAAAAGGGA GAGAGAGGIG AAGGCGGCTG TCCGCGCGAT CTGCTCTGCC 120
CGSACGCAGC TGTTACGACC TCTCAGTCT CAGTCAGGAG CAAG ATG GGT GCC GGC 176
Met Gly Ala Gly
1
GGC AGG ATG ACC GAG AAG GAG CGG GAG AAG CAG GAG CAG CTC GGC CGA 224
Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu Gln Leu Ala Arg
5 10 15 20
GCT ACC GGT GGC GGC GCG ATG CAG CGG TCG CCG GTG GAG AAG CCT CCG 272
Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val Glu Lys Pro Pro
25 30 35
TTT ACT CTC GGT CAG ATC AAG AAG GGC ATC CCG CCA CAG TGC TTC GAG 320
Phe Leu Ile Gly Thr Ile Lys Lys Ala Ile Ile Ile His Lys Thr Glu
40 45 50
CGG TCG GTG CTC AAG TCG TTC TCG TAC GTG GTC CAC GAC CTG GTG ATC 368
Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile
55 60 65
GGG GCG GCG CTC CTC TAC TTC GCG CTG GGC ATC ATA CCG GCG CTC CGA 416
Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile Pro Ala Leu Pro
70 75 80
AGC CCG CTC CGC TAC GGT GCG TGG CCG CTG TAC TGG ATC GCG CAG GGG 464
Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp Ile Ala Gln Gly
85 90 95 100
TGT GTG TGC ACT GGT GTG TGG CTC ATC GCG TAC GAG TGC GGC CAC CAC 512
Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His
105 110 115
GGG TTC TCG CAG TAC TCG CTC CTC CAG CAG CTC CTC GAG CTC GAG CAG 560
Ala Gln Ser Arg Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
120 125 130
CAC TCG TCG CTC ATG GTG CCG TAC TTC TCG TGG AAG TAC AGC CAC CGG 608
His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg
135 140 145
CGC CAC CAC TCC AAC ACG GGG TCC CTG GAG CGC GAC GAG GTG TTC GTG 656
Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val
150 155 160
TTC AAC AAG AA: GAG GCG CTC CCG TCG TAC ACC CCG TAC GTG TAC AAC 704
```

Pro	Lys	Lys	Lys	Glu	Ala	Leu	Pro	Trp	Tyr	Thr	Pro	Tyr	Val	Tyr	Asn	
165					170					175					180	
AAC	CCG	GTC	GGC	CCG	GTG	GTG	CAC	ATC	GTG	GTG	CAG	CTC	AAC	CTC	GGG	752
Asn	Pro	Val	Gly	Arg	Val	Val	His	Ile	Val	Val	Gln	Leu	Thr	Leu	Gly	
			185					190						195		
TGG	CCG	CTG	TAC	CTG	GCG	ACC	AAC	GCG	TGG	GCG	CGG	CCG	TAC	CCG	CGC	800
Trp	Pro	Leu	Tyr	Leu	Ala	Thr	Asn	Ala	Ser	Gly	Arg	Pro	Tyr	Pro	Arg	
			200					205						210		
TTC	GCC	TGC	CAC	TTC	GAC	CCC	TAC	GGC	CCC	ATC	TAC	AAC	GAC	CGG	GAG	848
Phe	Ala	Cys	His	Phe	Asp	Pro	Tyr	Gly	Pro	Ile	Tyr	Asn	Asp	Arg	Glu	
		215					220						225			
CGC	GCC	CAG	ATC	TTC	GTC	TGC	GAC	GCC	GGC	GTC	GTC	GCC	GTG	GCG	TTC	896
Arg	Ala	Gln	Ile	Phe	Val	Ser	Asp	Ala	Gly	Val	Val	Ala	Val	Ala	Phe	
	230					235						240				
GGG	CTG	TAC	AAG	CTG	GCG	GCG	GCG	TGC	GGG	GTC	TGG	TGG	GTG	GTG	CGC	944
Gly	Leu	Tyr	Lys	Leu	Ala	Ala	Ala	Phe	Gly	Val	Trp	Trp	Val	Val	Arg	
245					250					255					260	
GTG	TAC	GCC	GTG	CCG	CTG	CTG	ATC	GTG	AAC	GCG	TGG	CTG	GTG	CTC	ATC	992
Val	Tyr	Ala	Val	Pro	Leu	Leu	Ile	Val	Asn	Ala	Trp	Leu	Val	Leu	Ile	
				265				270						275		
ACC	TAC	CTG	CAG	CAC	ACC	ACC	CCG	TTC	CTC	CTC	CAC	TTC	GAC	TCC	ACC	1040
Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ser	Leu	Pro	His	Tyr	Asp	Ser	Ser	
			280					285						290		
GAG	TGC	GAC	TGG	CTG	CGC	GGC	GGC	CTG	GCC	ACC	ATG	GAC	CGC	GAC	TAC	1088
Glu	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Leu	Ala	Thr	Met	Asp	Arg	Asp	Tyr	
		295					300					305				
GGC	AIC	CTC	AAC	CGC	GTG	TTC	CAC	AAC	ATC	ACG	GAC	ACG	CAC	GTG	GCG	1136
Gly	Ile	Leu	Asn	Arg	Val	Phe	His	Asn	Ile	Thr	Asp	Thr	His	Val	Ala	
	310					315					320					
CAC	CAC	CTC	TTC	TGC	ACC	ATG	CGC	CAC	TAC	CAC	GCC	ATG	GAG	GCC	ACC	1184
His	His	Leu	Phe	Ser	Thr	Met	Pro	His	Tyr	His	Ala	Met	Glu	Ala	Thr	
	325					330					335				340	
AAG	GCG	ATC	AGC	CCC	ATC	CTG	GGC	GAC	TAC	TAC	CAC	TTC	GAC	CGG	ACC	1232
Lys	Ala	Ile	Arg	Pro	Ile	Leu	Gly	Asp	Tyr	Tyr	His	Phe	Asp	Pro	Thr	
				345										350		
CGT	GTC	GCG	AAG	GGG	ACC	GTC	GGC	GAC	TTC	TTC	GTC	GTC	GTC	GTC	GTC	
	355															
GAG	CGC	GAG	GAC	CGC	AAG	GGC	GTG	TTC	TGG	TAC	AAC	AAG	AAG	TTC	TACCCCGCCG	
1335																
Glu	Pro	Glu	Asp	Arg	Lys	Gly	Val	Phe	Trp	Tyr	Asn	Lys	Lys	Phe		
		370					380					385				
CGCTCGCAGA	GCTGAGBACC	CTACCGTAGG	AATBEGAGCA	BAAACCAAGA	GGAGGAGACG											1395
GTACTCGGCC	CAAAGTCTGC	GTAAAGCTAT	CTAATCGTTA	GTGCTGAGTC	TTTLAGAGGG											1415

GAAGAGAGAT CATTGCGGCA CAGAGAAGAA GGCTTACTGC AGTGCCATCG CTAGAGCTGC 1515
CATCAASTAC AAGTAGGJAA ATTCGTCAAC TTAGTGTGTC CCATGTTGTT TTTCTTAGTC 1575
GTCCGCTGCT GTAGGCTTTC CGGCGGCGGT GGTGTCGTC GTTGGCATCC GTGGCCATGC 1635
CTGTGCTGTC GTGGCCGCGC TTGTGCTGTC CGTCTGTGTC CGCGTTGGCG TGTCTCTTTC 1695
GTGCTCCCCG TGTGTTGTTG TAAAACAAGA AGATGTTTTC TGGTGTCTTT GGCGAATAA 1755
CAGATCGTCC GAAAGAAAAA AAAAAAAAAA AAAAA 1790

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 387 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
1 10 15
Gln Ser Ala Ser Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Phe Val
20 25 30
Glu Lys Pro Phe Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
35 40 45
His Cys Phe Glu Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His
50 55 60
Asp Leu Val Ile Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile
65 70 75 80
Pro Ala Leu Phe Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp
85 90 95
Ile Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu
100 105 110
Cys Gly His His Ala Phe Ser Arg Lys Ser Leu Ile Asp Asp Val Val
115 120 125
Tyr Leu Val Leu Ala Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys
130 135 140
Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
145 150 155 160
Glu Val Phe Val Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro
165 170 175
Tyr Val Tyr Asn Asn Phe Val Ser Arg Val Val His Ile Val Val Val
180 185 190

Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg
 195 200 205
 Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
 210 215 220
 Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val
 225 230 235 240
 Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp
 245 250 255
 Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp
 260 265 270
 Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His
 275 280 285
 Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
 290 295 300
 Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
 305 310 315 320
 Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
 325 330 335
 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His
 340 345 350
 Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu
 355 360 365
 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn
 370 375 380
 Lys Lys Phe
 385

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 673 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: 2 3 1
- (D) TOPOLOGY: linear

(ii) SEQUENCE TYPE: cDNA

(iii) SYNTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Ricinus communis

(vii) IMMEDIATE SOURCE:

(B) CLONE: pRF2-1C

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..673

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGG GTG ATG GCG CAT GAT DGT GGG CAC CAT GGC TTC AGT GAC TAT CAA	48
Trp Val Met Ala His Asp Cys Gly His Ala Phe Ser Asp Tyr Gln	
1 5 10 15	
TTG CTT GAT GAT GTA GTT GGT CTT ATC CTA CAC TCC TGT CTC CTT GTC	96
Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val	
20 25 30	
CCT TAT TTT TCA TGG AAA CAC AGC CAT GGC CGA CAT CAT TCC AAC ACA	144
Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His Ser Asn Thr	
35 40 45	
GGG TCC CTG GAA CGG GAT GAA GTG TTT GTT CCC AAG AAG AAA TCT AGT	192
Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser	
50 55 60	
ATC CGT TGG TAT TCC AAA TAC CTC AAC AAC CCT CCA GGT CGT ATC ATG	240
Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met	
65 70 75 80	
ACA ATT GGC GTG ACA CTT TCA CTT GGC TGG CCT CTG TAC CTA GCA TTC	288
Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe	
85 90 95	
AAT GTT TCA GGC AGG CCA TAT GAT GGG TTC GGC TGG CAC TAT GAC CCA	336
Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro	
100 105 110	
TAT GGC CCG ATC TAC AAT GAT CCG GAG CGA ATC GAG ATA TTC ATA TCA	384
Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser	
115 120 125	
GAT GCT GGT GTT CTT GCT GTC ACT TTT GGT CTC TAC CAA CTT GCT AIA	432
Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile	
130 135 140	
GGG AAG GGG CTT GGT TGG GTT GTC TGT GTA TAT GGA GTG CCA TTG TTG	480
Glu Lys Glu Ile Ala Trp Val Leu Tyr Val Tyr Gly Val Ile Leu Leu	
145 150 155 160	
CTG CTG AAT TCA TCG GTG GTT CCG AAT ACA TTT CCG CAG CAT ACT CAC	528
Val Val Asn Ser Phe Leu Val Thr Thr Phe Cys Val Ile Ile Ile Ile	
165 170 175	
CCT GCA TTG CCA CAT TAT GAT TCG TCG GAG TCG GAC TGG CTA AGA GGA	576
Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly	
180 185 190	
CTT CTA GCA ACT GTT GAT AAT GAT TAT GGG ATC TTG AAC AAG CTA TTC	624

Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
 195 200 205

CAT AAC ATA ACG GAC ACT CAA GTA GCT CAC CAG CTT TTC ACC ATG CCC C 673
 His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro
 210 215 220

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 224 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln
 1 5 10 15

Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val
 20 25 30

Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr
 35 40 45

Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser
 50 55 60

Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met
 65 70 75 80

Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe
 85 90 95

Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro
 100 105 110

Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser
 115 120 125

Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile
 130 135 140

Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu
 145 150 155 160

Val Val Ala Ser Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
 165 170 175

Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly
 180 185 190

Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
 195 200 205

His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Ile Thr Met Pro
 210 215 220

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1369 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) ORIGINAL SOURCE:

(A) ORGANISM: *Epizootic hemorrhagic septicemia virus*

(vii) IMMEDIATE SOURCE:

(B) CLONE: pRF1975-42

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 184..1347

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

CGCCCCGGGAT TTAGAGTTTC ACACCAATTT GCAAAAATG CAIGATTTCG CTTAAAATCA      60
AACACCACAC CTTATAACTT AGTCITAAGA GAGAGAGAGA GAGGAGACAT TTCTCTTCTC      120
TGAGATGAGC AATTCTCTTC CAGATATCGA AGCCTCAGGA AAGTGCTTGA GAAGAGCTTG      180
AGA ATG GGA GGT GGT GGT CGC ATG TCT ACT GTC ATA ATC AGC AAC AAC      228
Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn
   1           5           10           15
AGT GAG AAG AAA GGA GGA AGC AGC CAC CTG GAG CGA GCG CCG CAC ACG      276
Ser Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr
           20           25           30
AAG CCT CCT TAC ACA CTT GGT AAC CTC AAG AGA GGC ATC CCA CCC CAT      324
Lys Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His
           35           40           45
TGC TTT GAA GGC TCT TTT GTS CGC TCA TTC TCC AAT TTT GCC TAT AAT      372
Cys Phe Thr Ser Ser Ser Arg Ser Thr Thr Asn Thr Ala Tyr Asn
           50           55           60
TTC TGC TTA AGT TTT GTC TCC TAC TCG ATC GGC ACC AAC TTC TTC CCG      420
Phe Cys Thr Phe Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
           65           70           75
TAC ATC TCT TTT CCG GTC TCG TAT GGT GGT TGG GTC GTC TAC TCG CTC      468
Tyr Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu
           80           85           90           95
TTT TAA TTT TTT ATT TTT ATT TTT CTT TTT TTT ATT TTT TAT TAA TTT      516

```

Phe	Gln	Gly	Cys	Ile	Leu	Thr	Gly	Leu	Trp	Val	Ile	Gly	His	Glu	Cys	
				100					105					110		
GGC	CAT	CAT	GCT	TTT	AGT	GAG	TAT	CAG	CTG	GCT	CAT	CAC	ATT	GTT	GGC	564
Gly	His	His	Ala	Phe	Ser	Glu	Tyr	Gln	Leu	Ala	Asp	Asp	Ile	Val	Gly	
			115					110					125			
CTA	ATT	GTC	CAT	TCT	GCA	CTT	CTG	GTT	CCA	TAT	TTT	TCA	TGG	AAA	TAT	612
Leu	Ile	Val	His	Ser	Ala	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	
		131					135					145				
AGC	CAT	CGC	CGC	CAC	CAT	TCT	AAC	ATA	GGA	TCT	CTC	GAG	CGA	GAC	GAA	660
Ser	His	Arg	Arg	His	His	Ser	Asn	Ile	Gly	Ser	Leu	Glu	Arg	Asp	Glu	
	145					150					155					
GTG	TTC	GTC	CGG	AAA	TCA	AAG	TGG	AAA	ATT	TCA	TGG	TAT	TCT	AAG	TAC	708
Val	Phe	Val	Pro	Lys	Ser	Lys	Ser	Lys	Ile	Ser	Trp	Tyr	Ser	Lys	Tyr	
160					165				170					175		
TTA	AAC	AAC	CGG	CCA	GCT	CGA	GTT	TTG	ACA	CTT	GCT	GCC	ACG	CTC	CTC	756
Leu	Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu	
			180					185						190		
CTT	GGC	TGG	CCT	TTA	TAT	TTA	GCT	TTC	AAT	GTC	TCT	GGT	AGA	CCT	TAC	804
Leu	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	
			195				200						205			
GAT	CGC	TTT	GCT	TGC	CAT	TAT	GAT	CGC	TAT	GGC	CCA	ATA	TTT	TCT	GAA	852
Asp	Arg	Pro	Ala	Cys	His	Tyr	Val	Ala	Tyr	Gly	Ile	Ile	Pro	Ser	Glu	
		210					215					220				
AGA	GAA	AGG	CTT	CAG	ATT	TAC	ATT	GCT	GAC	CTC	GGA	ATC	TTT	GCC	ACA	900
Arg	Gln	Arg	Leu	Gln	Ile	Tyr	Ile	Ala	Asp	Leu	Gly	Ile	Phe	Ala	Thr	
	225					230					235					
ACG	TTT	GTG	CTT	TAT	CAG	GCT	ACA	ATG	GCA	AAA	GGG	TTG	GCT	TGG	GTA	948
Thr	Phe	Val	Leu	Tyr	Gln	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val	
240					245				250					255		
ATG	CGT	ATC	TAT	GGG	GTG	CCA	TTG	CTT	ATT	GTT	AAC	TGT	TTC	CTT	GTT	996
Met	Arg	Ile	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Cys	Phe	Leu	Val	
			260					265					270			
ATG	ATC	ACA	TAC	TTG	CAG	CAC	ACT	CAC	CCA	GCT	ATT	CCA	CGC	TAT	GGC	1044
Met	Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly	
			275					280					285			
TCA	TGG	GAA	TGG	GAT	TGC	CGC	CGG	CCA	CCA	ATG	GTC	ACT	GTC	GAT	AGA	1092
Ser	Ser	Gln	Leu	Asp	Phe	Ala	Arg	Gln	Ala	Met	Val	Ile	Val	Val	Leu	
		290					295					300				
GAT	TAT	GGG	GTG	TTG	AAT	AAA	GGA	TTT	CAT	AAC	ATT	GCA	CAC	ATT	TAT	1140
Asp	Tyr	Gly	Val	Leu	Asn	Lys	Val	Phe	His	Asn	Ile	Ala	Asp	Thr	His	
	305					310					315					
GTA	GCT	CAT	CAT	CTC	TTT	GCT	ACA	GTG	CCA	CAT	TAC	CAT	GCA	ATG	GAG	1188
Val	Ala	His	His	Leu	Phe	Ala	Thr	Val	Pro	His	Tyr	His	Ala	Met	Glu	
320					325				330					335		

GCC ACT AAA GCA ATC AAG CCT ATA ATG GGT GAG TAT TAC CGG TAT GAT	1236
Ala Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp	
340 345 350	
GGT ACC CCA TTT TAC AAG GCA TTG TGG AGG GAG GCA AAG GAG TGC TTG	1284
Gly Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu	
355 360 365	
TTC GTC GAG CCA GAT GAA GGA GCT CCT ACA CAA GGC GTT TTC TGG TAC	1332
Phe Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr	
370 375 380	
CGG AAC AAG TAT TAAAAAGTG TCATGTAGCC TGCCG	1360
Arg Asn Lys Tyr	
385	

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn Ser	1 10
Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr Lys	20 30
Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His Cys	35 45
Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Asn Phe Ala Tyr Asn Phe	50 60
Cys Leu Ser Phe Leu Ser Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr	65 75 80
Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe	85 90 95
Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly	100 105 110
Asn Glu Ala Ile Ser Val Tyr Val Leu Ala Asp Asp Ile Val Gly Leu	115 120 125
Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Thr Ser	130 135 140
His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val	145 150 155 160
Phe Val Pro Lys Ser Lys Ser Lys Ile Ser Thr Tyr Ser Lys Tyr Leu	165 170 175

Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu	Leu
			180					185					190		
Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp
		195					200					205			
Arg	Phe	Ala	Cys	His	Tyr	Asp	Pro	Tyr	Gly	Pro	Ile	Phe	Ser	Glu	Arg
	210					215					220				
Glu	Arg	Leu	Gln	Ile	Tyr	Ile	Ala	Asp	Leu	Gly	Ile	Phe	Ala	Thr	Thr
	225				230					235					240
Phe	Val	Leu	Tyr	Gln	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val	Met
				245					250					255	
Arg	Ile	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Cys	Phe	Leu	Val	Met
			260					265					270		
Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly	Ser
		275					280					285			
Ser	Glu	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Met	Val	Thr	Val	Asp	Arg	Asp
	290					295					300				
Tyr	Gly	Val	Leu	Asn	Lys	Val	Phe	His	Asn	Ile	Ala	Asp	Thr	His	Val
	305				310					315					320
Ala	His	Leu	Ser	Phe	Ala	Ile	Val	Ile	His	Lys	His	Ala	Met	Glu	Ala
				325					330					335	
Thr	Lys	Ala	Ile	Lys	Pro	Ile	Met	Gly	Glu	Tyr	Tyr	Arg	Tyr	Asp	Gly
			340					345					350		
Thr	Pro	Phe	Tyr	Lys	Ala	Leu	Trp	Arg	Glu	Ala	Lys	Glu	Cys	Leu	Phe
		355					360					365			
Val	Glu	Pro	Asp	Glu	Gly	Ala	Pro	Thr	Gln	Gly	Val	Phe	Trp	Tyr	Arg
	370					375						380			
Asn	Lys	Tyr													
	385														

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(a) LENGTH: 23 base pairs

(b) TYPE: coding

(c) STRANDEDNESS: single

(d) ORIENTATION: 5' to 3'

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 1.13
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGGTATGCC ATGANTGTGG NCA

23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION: 1.12
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAATGTGTG CACGAGGTG TC

22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1973 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) ORIGINAL SOURCE:

HEP-2 CELL LINE (ATCC CCL-22) (American

type culture collection)

(vi) IMMEDIATE CLONE:

(B) CLONE: pAGE2-6

(ix) FEATURE:

(A) NAME/KEY: exon
(B) LOCATION: 433..526

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 521..1654

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

ATTGGGTAAT TCCACATAT TTTAGAGATT AGTTTSGAGT TCCATCCATA CTTTACTAGT    60
GATTATAAAT TTAAAATACG TACTTTTGGG CTATAAAAGT AAACTAAGTA AATTAGAACG    120
TGATATTAAA AAGTTAATGT TCACTGTTAT ATTTTTTTCA CAAGTAAAAA ATGGGTTATT    180
TGGGGTAAAT AAAAAATCCA GATATTTTGA ATTGATTAAA AAGGTTGAAA TAAGAGAGGA    240
GGGGAAAGAA AAGAAGGTGG GGGGCCASTA TGAAAAGGGA AGGTGTGATC AAATCATCTC    300
TCTCTCTCTC TACCTTGGAC CCACGGGCGG TGTCCATTTA AAGGCCGTGC TCTTGGCATI    360
CCCCATCTGA CCATCAGAAG AAGAGCCACA CACTCACAAA TTAAAAAGAG AGAGAGAGAG    420
AGAGAGACAG AGAGAGAGAG AGATTCTGCG GAGGAGCTTC TTCTGCTAG GGTGTTCATC    480
GTTATTAAAG TTATGGCCCC TAGGTGAGCT CCATCTGAG GTCCTGCTCT TCTCTTCCAT    540
TTCTTCTCAT TTTGATTTT GATTCTTATT TCTTTCCAGT AGCTGCTGCT CTGTGAATTI    600
CTCGGCTCAC GATACATCTC CTTATCTCC TTACATTCAA CTTTAGATCT GGTCTGAT    660
CTCTCTCTCT TTTCTTTCTT CTTTCTCTCT AGAGCTGAT GTTCTTTAT GTTCTGCTAC    720
CATTAATAAT GATGAATCTT CCAATTCATA CAATGATTAG TTTCTCTGCT CTACCAAAAG    780

ATATGTTGCA TTTTCACTTT TCTTCTTTTT TTCTAAGATG ATTTGCTTTG ACCAATTTGT    840
TTAGATCTTT ATTTTATTTT ATTTTCTGGT GGGTTGGTGG AAATGAAAAA AAAAAAAAAA    900
AAAAGCATAA ATTGTTATTT GTTAATGTAT TCATTTTTTG GCTATTGTGT CTGGGTAAAA    960
ATCTGCTTCT ACTGTTGAAT CTTTCTCTGA TTTTCTACTC CTATTGGGTT TTTATAGTAA   1020
AAATACATAA TAAAAGGAAA AAAAAAGTTT TATAGATTCT CTTAAACCCG TTACGATAAA   1080
AGTTGCAATC AAAATATTC AGGATCAGAT GGTCTTTGAT TGATTCAGAT GCGATTACAG   1140
TTGCATGAA AATTTTCTAG ATCGCTGCTC ACATTTTATT TTCTGTTTAA ATATCTAAAT   1200
CTGATATATC ATCTGCTGAA ATTTTCTCT CTTATATAT ATTTGATTC TTCTCTTTT   1260
GTTTCTCTCT TCACTTCTCT TTTCAATATC ATTTTCTGAT TCGATCTCTC AATTTTAAAT   1320
ACAAGCAAA TGAATGTTAA CACAAGCAAG AGATGTGACC TGCTTTATTA ACATGCTATT   1380
ACCTACTACT AGTCTATTC TCAACGCAAT CTTTTTTGTA TTTCTTACAT TATGCGGCTT   1440
CTCTACTCTT TATTCCTTTT GGTCCAGGCA TTTCTATTT GTGGCAATCC CTTTCAAAAC   1500
CTGATTTCCC ACTTTGGATC ATTTGCTCTA AGACTCTCTT GAATCGTTAC CACTTGTTC   1560
TTCTGATCTT TTTTTTTTTT AAAAAATC ATAAAAAGTAT TCCATACTCT TGAGTTTTCA   1620

```


GCTTGTTGAT TCTTTTGGTT TTGGTTTTCT GCAGAAACAT GGSTGCAGGT GGAAGAATGC 1680
CGTTTCTTAC TTCTTCCAAAG AAATCGGAAA CCGACACCCAC AAAGCTGTGT CCGTGCAGAG 1740
AACCGCCTTT CTGGGTGGGA GATCTGAAAG AAGCAATCCC GCGGCATTGT TTCAAACGCT 1800
CAATCCCTCG CTCTTTCTCC TACCTTATCA GTGACATCAT TATAGCCTCA TGCTTCTACT 1860
ACGTGCGCAC CAATTACTTC TCTCTCCTCC CTCAGCCTCT CTCTTACTTG GCTTGGCCAC 1920
TCTATTGGGC CTGTCAAGGC TGTGTCTTAA CTGGTATCTG GTCATAGCC CACGAATGCG 1980
GTCACCAAGC ATTCAAGGAC TACCAATGGC TGGATGACAC AGTGGGTCTT ATCTTCCATT 2040
CCTTCTCTCT CGTCCCTTAC TTCTCTGGA AGTATAGTCA TCGCCCTCAC CATTCCAACA 2100
CTGATCTCT CGAAAGAGAT GAAATATTTG TCCCAAAGCA GAAATCAGCA ATCAAGTGGT 2160
ACGCGAAATA CCTCAACAAC CCTCTTGGAC GCATCATGAT GTTAACCGTC CAGTTTGTCC 2220
TCGGGTGCCC CTGTACTTAA GCGTTTAAAG TCTGTGGCAG ACCGTATGAC GGGTTGGCTT 2280
GCCATTCTTT CCCCACGCT CCCATCTACA ATGACCGAGA ACCGCTCCAG ATATACCTCT 2340
CTGATGCGGG TATTCTAGCC GTGTGTTTTG GTCTTTAGCG TTACGCTGCT GCACAAGGGA 2400
TGGCTTGAT GATCTGCTT TACGGAGTAC CCGTCTGAT AGTGAATGCG TTCTCTCTCT 2460
TGATCACTTA CTTGCAAGC ACCTACCTCT CGTGGCTCA CTAGGATTC TACAGGTGGG 2520
ACTGGCTGAG GGGAGCTTTG GCTACCGTAG ACAGAGACTA CGGAATCTG AACAAAGTGT 2580
TCCACAACAT TACAGACACA CAGGTGGCTC ATCAGCTGTT CTGACAATG CCGCATTATA 2640
ACGCAATGGA AGCTACAAA GCGATAAAGC CAATTCTGGG AGACTATTAC CAGTTCTGAT 2700
GAACACCGTG GTATGTGGG ATGTATAGGG AGGCAAAGGA GTGTATCTAT GTAGAACGGG 2760
ACAGGGAAGG TGACAAAGAA GGTGTGACT GTTACAACAA TAAGTTATGA GGATGATGGT 2820
GAAGAAATG TCGACTTTTC TCTTGTCTGT TTGTCTTTTG TTAAAGAAAG TATGCTTCTG 2880
TTTAATAATC TTATTGTCCA TTTTGTGTG TTATGACATT TTGGGTGCTC ATTATGTTAT 2940
CTGTGA TTA TTAAGG TTTT TTTT 3000

(ii) INFORMATION FOR SEQ. 15 (M15):

(a) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) MOLECULE TYPE: cDNA

(iv) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION: 1..23
(D) OTHER INFORMATION: product=
"synthetic"
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGCATGTNG ARAANAERTG RTG

23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYBRIDICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION: 1..23
(D) OTHER INFORMATION: product=
"synthetic"
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGGCATGTTC TAAANAERTG RTG

23

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Ile Pro Arg His Cys

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Ile Pro Lys His Cys

1 6

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 4

(D) OTHER INFORMATION: Xaa = Pro or Lys

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Ile Pro Xaa His Cys

1 6

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 3

(D) OTHER INFORMATION: Xaa = Leu or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Trp Pro Xaa Tyr Irp

1 5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Trp Pro Leu Tyr Trp

1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 3
- (D) OTHER INFORMATION: Xaa = Leu or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Trp Pro Xaa Tyr Trp
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ala His Glu Cys Gly His
1 6

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Gly His Asp Cys Gly His
1 6

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 4
- (D) OTHER INFORMATION: Xaa = Ala or Gly

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 3
(D) OTHER INFORMATION: Xaa = Asp or Glu

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Xaa His Xaa Cys Gly His
1 5

(2) INFORMATION FOR SEQ ID NO:27:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Leu Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:28:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ile Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:29:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) REGION:

(A) NAME/KEY: unsure
(B) LOCATION: 1
(D) OTHER INFORMATION: Xaa = Leu or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Xaa Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Trp Lys Tyr Ser His Arg
1 5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Trp Arg Ile Ser His Arg
1 6

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 1
(C) OTHER INFORMATION: Xaa = Arg or Lys

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 1
(C) OTHER INFORMATION: Xaa = Ile or Tyr

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Trp Xaa Xaa Ser His Arg
1 6

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ser His Arg Arg His His
1 5

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Ser His Arg Thr His His
1 5

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 4
(C) OTHER INFORMATION: Xaa = Arg or Thr

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser His Arg Xaa His His
1 5

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

His Thr Thr Thr Glu
1 5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Val Thr Tyr Leu His
1 5

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 1

(D) OTHER INFORMATION: Xaa = Ile or Val

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 5

(D) OTHER INFORMATION: Xaa = Gln or His

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Xaa Thr Tyr Leu Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Leu Pro His Tyr
1

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Leu Pro Trp Tyr
1

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 1
- (D) OTHER INFORMATION: Xaa = His or Trp

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Leu Pro Xaa Tyr

1

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 1
- (D) OTHER INFORMATION: Xaa = Arg or Lys

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Trp Leu Xaa Gly Ala Leu

1

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Leu Trp Ala Gly Xaa Leu

1

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 1
(D) OTHER INFORMATION: Xaa = Trp or Tyr

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 3
(D) OTHER INFORMATION: Xaa = Arg or Lys

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 5
(D) OTHER INFORMATION: Xaa = Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Xaa Leu Xaa Gly Xaa Leu
1 4

(2) INFORMATION FOR SEQ ID NO:45:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Thr Val Asp Arg Asp Tyr Gly
1 5

(2) INFORMATION FOR SEQ ID NO:46:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Thr Leu Asp Arg Asp Tyr Gly
1 5

(2) INFORMATION FOR SEQ ID NO:47:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: unsure

(B) LOCATION: 2
(D) OTHER INFORMATION: Xaa = Leu or Val

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Thr Xaa Asp Arg Asp Tyr Gly
1 6

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Thr His Val Ala His His Leu Phe
1 6

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Thr His Val Ile His His Leu Phe
1 6

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(a) NAME: 100% identical

(b) LOCATION: 4

(c) OTHER INFORMATION: identical to sequence no. 100

(x) SEQUENCE ALIGNMENT: No. 100, 101, 102, 103

Thr His Val Xaa His His His Leu Phe
1 5

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

His His Leu Phe Ser Thr Met Pro His Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

His His Leu Phe Pro Gln Ile Pro His Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 4

(D) OTHER INFORMATION: Xaa = Pro or Ser

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 6

(D) OTHER INFORMATION: Xaa = Gln or Thr

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 7

(D) OTHER INFORMATION: Xaa = Ile or Met

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

His His Leu Phe Ser Thr Met Pro His Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs

(B) TYPE: nucleic acid

(C) STRANGLINEAR: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (ix) FEATURE:
 (A) NAME KEY: misc feature
 (B) LOCATION: 1..12
 (D) OTHER INFORMATION: /product=
 "synthetic
 oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

ATAGCCCCCC AA

12

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 12 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TGGTCTTTTG GT

12

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: cDNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (ix) FEATURE:
 (A) NAME KEY: misc feature
 (B) LOCATION: 1..15
 (D) OTHER INFORMATION: /product=
 "synthetic
 oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CTAGATATGG CG CC

14

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME KEY: misc_feature
- (B) LOCATION: 1..10
- (C) OTHER INFORMATION: /product="synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTCGACGAGG

10

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME KEY: misc_feature
- (B) LOCATION: 1..12
- (C) OTHER INFORMATION: /product="synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AGATCTGATA CT

12

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ala Ile Pro Pro His Cys Phe
1 5

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 3

(D) OTHER INFORMATION: Xaa = Leu or Pro

(ix) FEATURE:

(A) NAME/KEY: unsure

(B) LOCATION: 4

(D) OTHER INFORMATION: Xaa = Ile or Leu or Val

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Ala Trp Xaa Xaa Tyr Trp
1 6

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

His Glu Cys Gly His
1 5